Fast tracking of moving objects using single-pixel imaging

DONGFENG SHI,[1,4] KAIXIN YIN,[2] JIAN HUANG,[1, 3*] KEE YUAN,[1] WENYUE ZHU,[1] CHENBO XIE,[1] DONG LIU,[1] AND YINGJIAN WANG[1,3]

[1]*Key Laboratory of Atmospheric Optics, Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, 230031, China.*
[2]*Research Center for Laser Physics and Technology, Key Lab of Functional Crystal and Laser Technology, Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, Beijing, 100190, China.*
[3]*University of Science and Technology of China, Hefei, 230026, China.*
[4]*Key Laboratory of Optical Engineering, Chinese Academy of Sciences, Chengdu, 610209, China.*

**jhuang@aiofm.ac.cn*

Abstract: Successive images of a scene are captured and then further processed to achieve the moving object tracking. However, due to modulation rate limitations of the spatial light modulator in single-pixel imaging (SPI) system, the imaging frame rate cannot meet the high-resolution and real-time requirements for object tracking. In this paper, we demonstrate a fast object tracking technique based on SPI with an ultra-low sampling rate that is independent of imaging. We construct modulation information that satisfies the projection conditions and can transform 2D images into 1D projection curves. The 1D projection curves, which provide the location information of the moving object, can be obtained with high resolution in real-time, and then the tracking of the moving object is realized. A background subtraction technique for tracking moving objects that removes static components from a scene is also proposed. The proposed technique is verified by computational simulations and laboratory experiments. In the laboratory experiments, we demonstrate that the proposed method can be used to track moving objects with less than 0.2% of the measurements established by the Nyquist criterion, and it presents a resolution of 256×256 pixels at ~177 fps. The reported technique accelerates the tracking speed of SPI and provides an efficient strategy for remote sensing and biomedical applications.

Keywords: Computational imaging; Image reconstruction techniques; Target tracking.

1. Introduction

Image inputs are among the most important sources of information for human production and life. Image detection techniques are the predominant approaches to obtain images and image information. Single-pixel imaging (SPI) [1-33], which is also referred to as ghost imaging or computational imaging, is a new imaging detection technique that employs un-scanned single pixel detector to reproduce an image of the object. According to the different modulation modes, SPI can be divided into forward modulation [1-16] and backward modulation [17,18] modes. In the forward modulation mode, illumination structured light is created by the spatial-light-modulator (SLM) or light source array and employed to illuminate the object, and the reflected/transmitted light is collected by a single-pixel detector. In the backward modulation mode, the image of the object is sampled by the SLM, and the sampled information is detected by a single-pixel detector. SPI based on the backward modulation mode can work in either natural background light or active illumination conditions, although SPI based on the forward modulation mode can work only in active illumination conditions. Using the correlation between the detected intensities and the modulation information of the SLM or light source array, the image of the object can be recovered. Since SPI uses an un-spatially resolved detector to obtain spatially resolved information, it requires a large amount

of different modulation information from the SLM or light source array at different times. Therefore, SPI sacrifices temporal resolution in exchange for spatial resolution. Nevertheless, since SPI presents certain advantages, such as high SNR imaging under low light conditions [16], wide imaging spectra [17], it has received widespread attention in recent years.

Currently, many scholars have studied images of moving objects using SPI [19-23]. In addition to imaging, tracking the position of a moving object is of paramount importance for remote sensing and biomedical applications [34]. Tracking moving objects is commonly accomplished by RADAR and LADAR technologies, which rely on directional probe beams that can be scanned spatially or angularly. Methods of employing SPI for object tracking have also attracted the attention of researchers. Traditional tracking methods must obtain time series images, and then a follow-up processing algorithm is utilized to track the object [21-23]. In [21], Omar et al proposed a method of tracking an object with 2.44% of the number of measurements established by the Nyquist criterion, and the resolution was 64×64 pixels and entangled photons were used. In their setup, each scene reconstruction took 13.3 minutes due to the low photon flux, which limits the system's real-time ability to track moving objects. A method of compressive moving target tracking with thermal light based on complementary sampling was proposed, but it was not tested via real-time tracking experiments [22]. In [23], the authors acquired 32×32 pixel real-time video for three-dimensional object tracking at 14 frames-per-second (fps) using SPI. The performance of traditional tracking algorithms based on capturing successive images is closely related to the imaging speed. Thus, increasing the imaging speed of the SPI to image dynamic scenes can improve the tracking efficiency. Recently, the application of deep learning with convolutional auto-encoder networks was shown to recover real-time 128×128 pixel video at 30 fps [10]. The authors [11] demonstrate experimental single-pixel detection with real-time reconstruction obtained in parallel with measurement at a frame rate of 11 Hz for highly compressive measurements with a resolution of 256 × 256. Only recently, an SPI scheme using an LED-based high-speed illumination module was reported, and an exciting frame rate of 1000 fps with 32×32 pixel resolution was achieved [12]. This method was based on forward modulation mode and can work only in active illumination conditions. The imaging resolution is limited by the LED array, and achieving a resolution over the megapixel level (similar to that observed with a SLM, such as in digital micro-mirror devices (DMDs) that present a 2560×1600 pixel resolution) is difficult. Due to the inverse relationship between the imaging frame rate and the imaging pixel resolution, increasing the imaging frame rate must reduce the imaging pixel resolution. For example, under the same compression sampling rate, the imaging frame rate is 1000 fps with a 32×32 pixel resolution and 15.6 fps with a 256×256 pixel resolution.

In SPI, the DMD system, which has high reflectivity, high light throughput, high frame rates and high spatial resolution [35], is popularly used as a SLM. According to the reversibility of light, the SPI method based on DMD can work well under both active and passive conditions. Therefore, studying SPI based on DMD is of great significance. Due to the limitation of the DMD's modulation rate, the imaging frame rate of the technologies referenced above cannot meet the high-resolution and high real-time requirements of object tracking. However, to track a moving object, one doesn't necessarily need to 'see' it [34]! Capturing successive images of a scene for further processing is not critical for object tracking. Thus, we propose a novel effective method based on DMD-SPI that can track moving objects with high pixel resolution and an ultra-low sampling rate, and it does not depend on capturing successive images. In our real-time tracking experiments, we can achieve 256×256 pixels in real-time video at ~177 fps with less than 0.2% of the measurements established by the Nyquist criterion. The layout of the present paper is organized as follows. In Section 2, we introduce the principles and deduction methods. In Section 3, simulations and experiments are conducted to evaluate our proposed methods. In Section 4, the conclusions of this work are summarized.

2. Principles and methods

According to the principle of SPI, modulation information S_n is used to modulate the light in the n-th time and a clear pattern image correlated with S_n is presented for the scene that contains the moving objects. The reflected light from the scene is detected by a single-pixel detector, and the detection intensity I_n can be expressed as

$$I_n = \sum_{x,y} f_t(x,y) \cdot S_n(x,y). \tag{1}$$

where x and y are the space coordinates, f_t indicates the scene information, including the moving objects at time t, and $\sum_{x,y}$ is the summation operation along the x and y directions. The image of the scene can be recovered by performing a correlation operation using the detected intensity values and modulation information. As described above, due to the limitation of the modulation rate of the DMD, the imaging frame rate for SPI is still not high enough for imaging fast moving object with high resolution. In other words, the imaging frame rate cannot meet the high-resolution and high real-time requirements of object tracking. Currently, object tracking is a challenging problem for SPI using traditional methods based on capturing successive images.

However, to track a moving object, we need to obtain only the positional information of the object in real time. In the traditional strategy, the positional information of the object is obtained through processing the captured successive images. Positional information is the product of successive images. In other words, capturing successive images is not a necessary part of object tracking [34]. Due to the constraint of the imaging frame rate of the SPI, a novel technique must be adopted that is different from the traditional imaging strategy to recover the positional information of the object. Here, we propose a new method to achieve the positional information of the objects in the scene from the 1D projection curves, which enable us to confirm the location where the object is located in the scene based on the edge of the projection curves [9]. The projection curve of the $N\times M$-pixels image $f_t(x,y)$ onto the x-axis is expressed as $f_{t,y}(x)$, whereas the projection curve onto the y-axis is $f_{t,x}(y)$. The projection curves $f_{t,x}(y)$ and $f_{t,y}(x)$ can be expressed as

$$f_{t,x}(y) = \sum_x f_t(x,y), \tag{2}$$

$$f_{t,y}(x) = \sum_y f_t(x,y), \tag{3}$$

where $\sum_x$ and $\sum_y$ represent the integral operations along the x and y directions, respectively. Performing the same operation on the modulation information S_n, the equations can be transformed to

$$S_{x,n}(y) = \sum_x S_n(x,y), \tag{4}$$

$$S_{y,n}(x) = \sum_y S_n(x,y). \tag{5}$$

In these expressions, the projection curve of $S_n(x,y)$ onto the x-axis is expressed as $S_{x,n}(y)$, whereas the projection curve onto the y-axis is $S_{y,n}(x)$. Figure 1 presents a schematic diagram of obtaining the projection curves. Figures 1(A) and 1(B) present the projection curves of the scene and the modulation information, respectively.

The parameters are adjusted to satisfy the projection conditions.

$$I_{nx} = \sum_y \left[\sum_x f_t(x,y) \cdot S_{x,n}(x,y) \right] = \sum_y \left[f_{t,x}(y) \cdot S_{x,n}(y) \right] \Big/ M, \tag{6}$$

$$I_{ny} = \sum_x \left[\sum_y f_t(x,y) \cdot S_{y,n}(x,y) \right] = \sum_x \left[f_{t,y}(x) \cdot S_{y,n}(x) \right] \Big/ M. \tag{7}$$

The above formulas show that the intensities of the interactions between the illumination pattern and the scene coincide with the intensities of the interactions between the projection curves of the scene and the modulation information. A specific form of the modulation information S_n is designed to satisfy the above equations. Here, the projection modulation information Sx and Sy with mutually orthogonal properties is proposed and constructed from a Hadamard matrix. In Fig. 2, the 3rd order Hadamard matrix is taken as an example. Each row ($S_{x,1}$, $S_{x,2}$……$S_{x,8}$) of the Hadamard matrix is employed to constitute Sx, and each column ($S_{y,1}$, $S_{y,2}$……$S_{y,8}$) of the Hadamard matrix is employed to constitute Sy. When the modulation information $S_x(x,y)$ is constructed, each row in matrix $S_x(x,y)$ is equal to a row of data in Sx, and 8 modulation information matrices [$S_{x,1}(x,y)$, $S_{x,2}(x,y)$……$S_{x,8}(x,y)$] can be obtained. For example, each row of the matrix $S_{x,4}(x,y)$ is equal to the row $S_{x,4}$. When the modulation information $S_y(x,y)$ is constructed, each column in matrix $S_y(x,y)$ is equal to a column of data in Sy, and 8 modulation information matrices [$S_{y,1}(x,y)$, $S_{y,2}(x,y)$……$S_{y,8}(x,y)$] can be obtained. For example, each column of the matrix $S_{y,4}(x,y)$ is equal to the column $S_{y,4}$. The object is illuminated with modulated illumination patterns $S_x(x,y)$ and $S_y(x,y)$, and then the corresponding detected intensities I_{nx} and I_{ny} are obtained, respectively. Figure 2 shows that the modulation information matrixes $S_x(x,y)$ and $S_y(x,y)$ can satisfy the conditions in Eqs. (6) and (7).

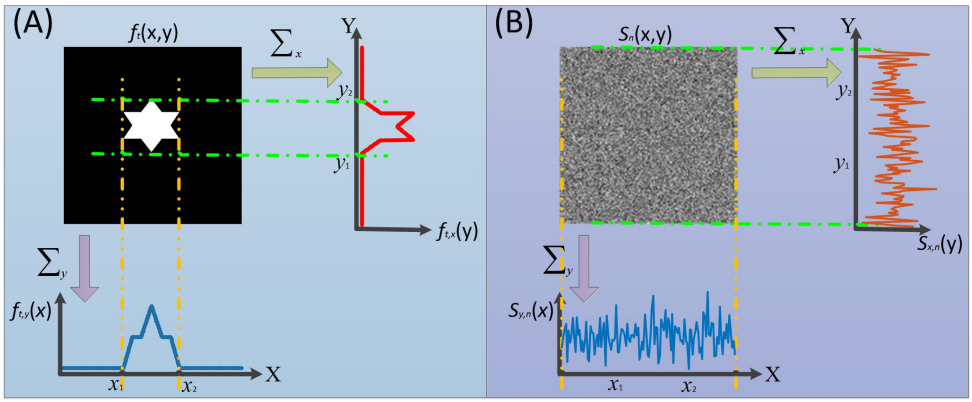

Fig. 1 Vertical and horizontal projection curves for A) scene with an object (star) on a uniform background and B) modulation information. The resolution of the image is 256×256 pixels.

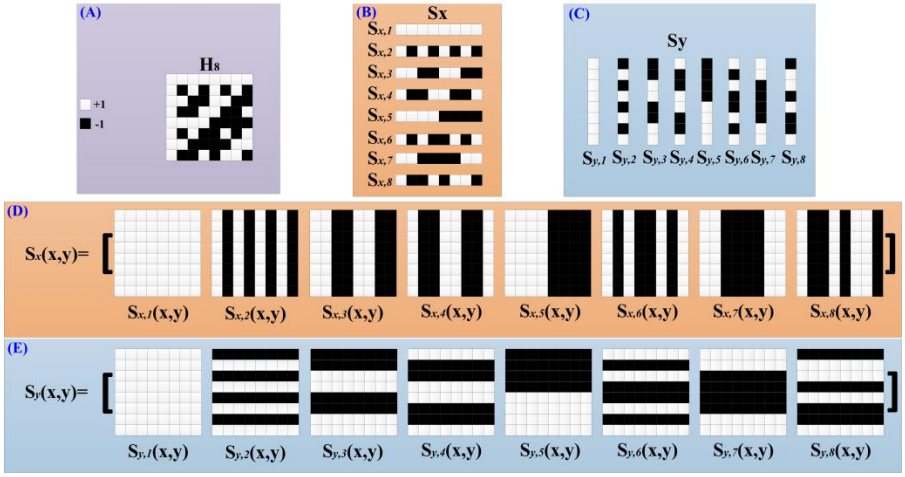

Fig. 2 Structure of the modulation projection matrix. (A) is the 3rd order Hadamard matrix. (B) Sx and (C) Sy are the projection modulation information. Modulated illumination patterns (D) $S_x(x,y)$ and (E) $S_y(x,y)$ are employed to illuminate the object.

Since the modulation information applied by the DMD is a previously known amount, the projection curves of the illumination pattern in both directions are also known. Finally, according to the principle of SPI, detected intensities combined with *Sx* and *Sy* are employed to recover the projection curves $f_{t,x}(y)$ and $f_{t,x}(y)$ of the objects, respectively. The formulas can be expressed as

$$f_{t,x}(y) = \sum_n \left[I_{nx} \cdot S_{x,n}(y) \right] / M , \tag{8}$$

$$f_{t,y}(x) = \sum_n \left[I_{ny} \cdot S_{y,n}(x) \right] / M . \tag{9}$$

In the above analysis, the background is assumed to be uniform. In many practical situations, the object will move in a complex background. In such cases, a background subtraction method [21] must be employed to remove background interference and obtain an accurate object motion trajectory. In this situation, the background will first illuminate with modulation information $S_x(x,y)$ and $S_y(x,y)$, and then the corresponding background reflected intensities I_{nxb} and I_{nyb} are acquired. When a moving object enters the scene, the corresponding detection intensities I_{nx} and I_{ny} can also be received. The formulas for obtaining the projection curves of the moving object in the complex background at time t are represented by

$$f_{t,x}(y) = \sum_n \left[(I_{nx} - I_{nxb}) \cdot S_{x,n}(y) \right] / M , \tag{10}$$

$$f_{t,y}(x) = \sum_n \left[(I_{ny} - I_{nyb}) \cdot S_{y,n}(x) \right] / M . \tag{11}$$

When the projection curves $f_{t,x}(y)$ and $f_{t,y}(x)$ of the object are recovered, a mutation phenomenon occurs on the positions where the object edges are located because the grayscale distribution differs between the object and the background. The edge detection algorithm [36] of first-derivative is employed to confirm the edges of the object region. The magnitude of first-derivative can be used to detect the presence of an edge at a point in an image. By acquiring the time series positional information of the moving object, tracking of the moving object can be achieved. Let the object region be $\Omega_t(x,y)$ at time t. Then, we have

$$\Omega_t(x,y), \quad \begin{cases} x_1 \le x \le x_2 \\ y_1 \le y \le y_2 \end{cases} , \tag{12}$$

where x_1 and x_2 represent the edges of the object region along the x-axis and y_1 and y_2 represent the edges of the object region along the y-axis, respectively. According to the positional information obtained at different times, continuous tracking of the moving object can be achieved. In our method, the recovered information is converted from 2D image information to 1D projection information, which greatly reduces the amount of information to be restored and improves the real-time object tracking performance. The next section will present experimental studies of the proposed method.

3. Experiments

3.1 Computational simulations

Computational simulations are employed to study the proposed method. The images for the experiments are shown in Fig. 3, where A represents the object to be tracked, B is a complex background scene, and C is the object in the complex scene. The size of the images is 256×256 pixels. An 8^{th} order Hadamard matrix is generated, and then the modulated projected patterns *Sx* and *Sy* in the x and y directions, respectively, are obtained according to the rules of the above section. Furthermore, Russian Dolls ordering [18] of the modulated

projected patterns is employed to compress the tracking. The compression sampling rate can be defined as

$$\gamma = K/M^2 . \tag{13}$$

In this equation, K is the number of modulated patterns and M^2 is the number of pixels in the image.

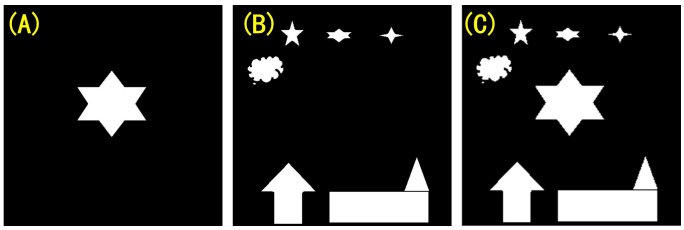

Fig. 3 Experimental scenes: A) object, B) complex background, and C) object in the complex background

The first experiment performed positioning of the object in a uniform black background as shown in Fig. 3A. The projected curves of the object under different numbers of modulated patterns are shown in Fig. 4. The illustration shows that as the number of samples increases, the recovered projection curves become more accurate. The percentage of root mean square error (RMSE) is used to quantify the restored projection curves, which can be expressed as

$$R(f_{x,k}) = \sqrt{(f_{x,k}(y) - f_x(y))^2 / M}, \tag{14}$$

$$R(f_{y,k}) = \sqrt{(f_{y,k}(x) - f_y(x))^2 / M}. \tag{15}$$

where $R(f_{x,k})$ and $R(f_{y,k})$ represent the RMSEs of the recovered projection curves; $f_{x,k}$ and $f_{y,k}$ represent the recovered projection curves in the y and x directions when the sampled number is k, respectively; and f_x and f_y represent the real projection curves in the y and x directions, respectively. The results are shown in Fig. 5, and they indicate that the recovered curves are closer to the true projection curves as the number of samples increases. When the number of samples is greater than 32, the RMSE is smaller than 10%. However, when tracking a moving object, region information $\Omega_t(x,y)$ of the object must be obtained. Next, edge detection of the projection curves is employed to acquire the parameters x_1, x_2, y_1 and y_2. The bright points in Fig. 6 show the corresponding parameters x_1, x_2, y_1 and y_2 under different sampling numbers. It can be seen that there is an error in the obtained position parameters of the object when the number of samples is small. However, as the number of samples increases, the acquired positional parameters gradually become stable. When the number of projections in the x and y directions is greater than 128, the acquired positional parameters do not change. In other words, when the sampling number is larger than a certain value, the obtained positional parameters and the accuracy of the tracking are not changed.

The second experiment performed positioning of the object in the complex scene as shown in Fig. 3(C). First, the complex scene is illuminated, and the reflected intensities I_{nxb} and I_{nyb} are obtained. Then, the scene that contains the object is illuminated, and the corresponding reflected intensities I_{nx} and I_{ny} are received. The results under different numbers of samples with Eqs. (10) and (11) used to restore the object's projection curves are shown in Fig. 7. The results show that the background subtraction technique can remove the static components from a scene and accurately obtain the positional parameters of the object.

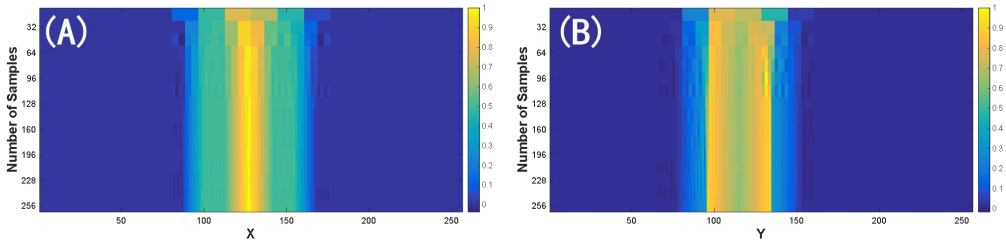

Fig. 4 Projection curves for the figure 3(A) scene with an object (star) on a uniform background under different sampling numbers: A) x-direction projection distribution and B) y-direction projection distribution. The ordinate indicates the number of samples.

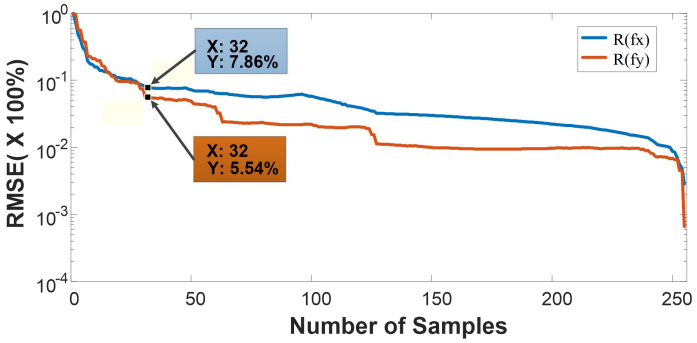

Fig. 5 RMSE vs. number of samples

Because Russian Dolls ordering is employed to sort the patterns, the resolution of the recovered information gradually increases as the number of samples increases. This rule can also be found from the results of Figs. 4 and 7. When the number of samples is small, the resolution of the restored projection curve is low. As the number of samples increases, the resolution of the projection curve becomes higher, and closer to the true value. However, because most energy of the spatial information of the projection curve can be recovered by utilizing the top-ranked patterns [18], we can simply acquire the most energy of components to reconstruct a high-quality projection curve. The fewer the number of patterns utilized, the higher fps achieved. When the requirement for positioning resolution is not high, this method can effectively increase fps.

The simulation results for the static object show that the proposed method can effectively achieve the positioning of the object. Due to the application of binary modulation of patterns in the proposed method, the high-speed binary modulation of DMD can be employed to achieve real-time tracking of moving objects with high resolution. In the next laboratory experiments, real-time tracking of moving objects will be studied.

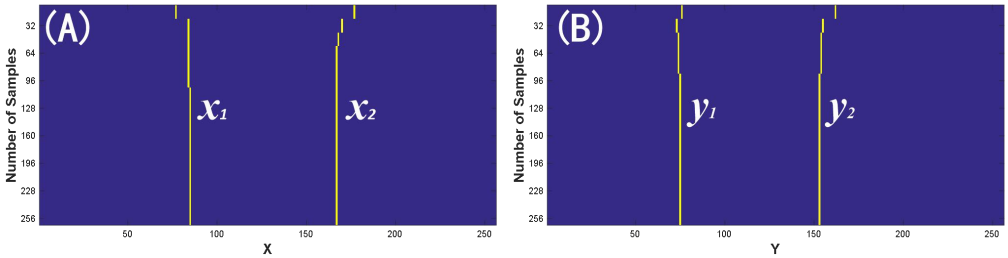

Fig. 6 Regional parameter results for the figure 3(A) scene with an object (star) on a uniform background. A) x_1 and x_2 represent the edges of the object region along the x-axis, and B) y_1 and y_2 represent the edges of the object region along the y-axis.

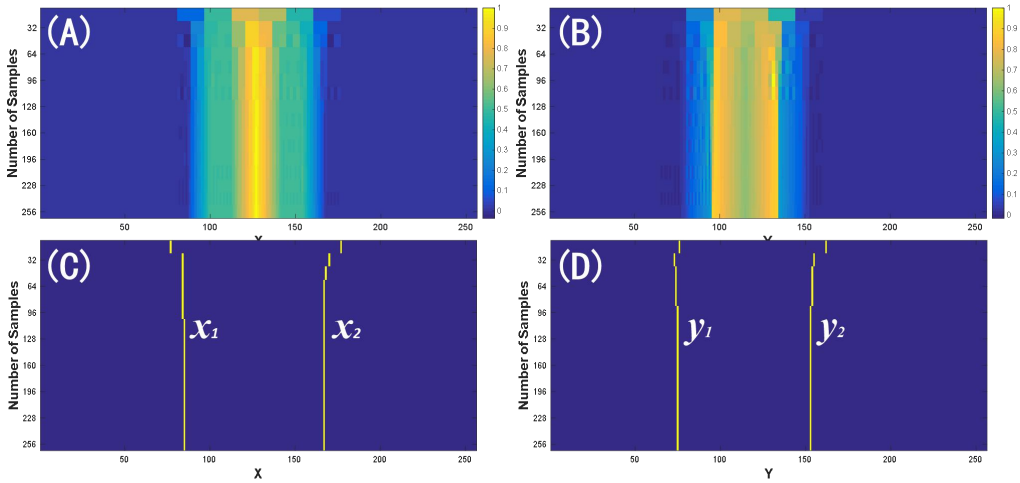

Fig. 7 Positioning results for the figure 3(C) scene with an object (star) on a complex background. A) x-direction projection distribution and B) y-direction projection distribution, C) x_1 and x_2 represent the edges of the object region along the x-axis, and D) y_1 and y_2 represent the edges of the object region along the y-axis.

3.2 Laboratory experiments

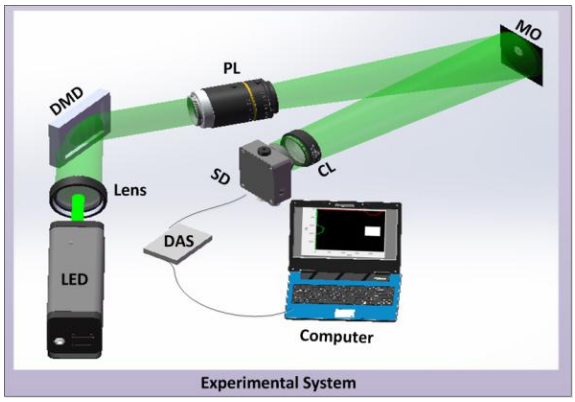

Fig. 8. Experimental system. MO is moving object. SD is single-pixel detector. DAS is data acquisition system. PL is projection lens group composed of five lenses. CL is collecting lens.

The proposed technique is studied using an experimental system described as follows. A 10 W white LED serves as the light source. A DMD system (Texas Instruments Discovery V7100 with 1024×768 micro-mirrors) is used to generate the illumination patterns. A single-pixel detector (SD, Thorlabs PMT-PMM02) and data acquisition system (Pico 6407 with sampling rate 200 MS/s) are employed for light detection and data acquisition, respectively. The light enters a lens and then reflected by the DMD, which provides the 2D patterns that are projected to the object by a projection lens (PL) with a focal length of 125 mm. The reflected light from the object is collected by a collecting lens (CL) with a focal length of 100 mm and then detected by the SD single-pixel detector. The computer used for running the experiment and processing the data is a National Instruments PXI system having an Intel Core i7 processor, RAM 12GB and running Windows 7. Next, the values of the intensities are sent to a computer via the data acquisition system (DAS). The 3×3 mirrors of the DMD are combined into a pattern cell that corresponds to an image pixel, and the intermediate 768×768 mirrors are utilized in the experiment. Thus, the resolution of the image is 256×256 pixels. The object is a white button with a diameter of ~3cm suspended with a black string, and the button is moved by pulling the string. The experimental setup is shown in Fig. 8.

A Hadamard pattern is constructed with values of either +1 or −1. Because the illumination patterns produced by the DMD system are binary, the positive and negative reflection values cannot be readily utilized. To address this issue, the approach involves a pair of matrices that are related to the matrix by a subtraction operation. The details can be found in [33]. This processing can effectively deduct the influence of the background light. For the tradeoff between the real-time nature and the tracking accuracy, the Russian dolls ordering method [18] is used to select 128 patterns which are composed of 64 positive and 64 negative patterns for illumination and implement the compressed tracking experiment. The sampling rate is approximately 0.195%. The numbers of the modulation information $S_x(x,y)$ and $S_y(x,y)$ are 64. The DMD's projection frequency is set to 22.7 kHz; thus, the obtained frequency of the object's position is approximately 177 fps.

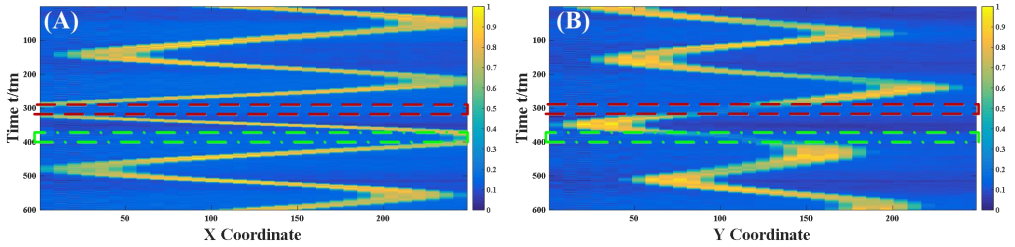

Fig. 9. Experimental results of the projection curves at different times when the white button randomly moves in a uniform black background: A) projection curves of the X coordinate and B) projection curves of the Y coordinate.

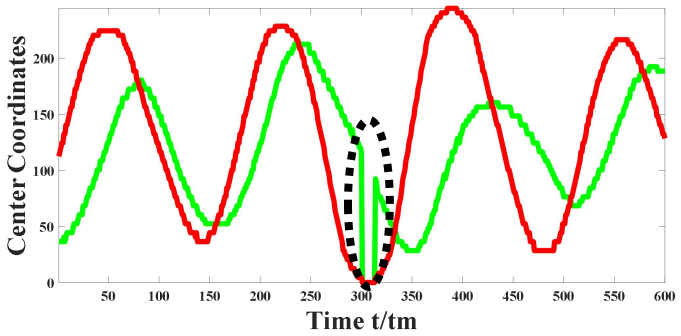

Fig. 10. Center distribution curve. The red line is the x-axis coordinate, and the green line is the y-axis coordinate. The algorithm automatically sets the position of the object to 0, when the object is out of the field of view. The black oval dashed line indicates that the object moves out of the field of view. The parameter tm denotes one measurement duration, which is equal to approximately 5.6 milliseconds.

First, the white button randomly moves in a uniform black background. The moving object is located at a distance of ~1.6m from the imaging system, and size of the illuminated area by the DMD is approximately 13.5×13.5cm at the object space. The proposed method is employed to trace the moving button. We conduct 600 sets of consecutive projections and probe acquisitions. The real-time acquisition of the object's position is achieved using the acquired data. The experimental results are shown in Fig. 9, where panel A shows the *x*-direction projection distributions of the object at different times and panel B shows the *y*-direction projection distributions of the object at different times. The ordinate is the time axis, and the data for each row represent the projection curve at different times. During the experiment, an object is moved from the field of view (FOV) for a period of time, and this process is accurately recorded and marked with a red square in Fig. 9. During this period of time, the projection distribution does not have an apparent edge compared with that of the object in the FOV. At this moment, the algorithm automatically sets the position of the object to 0. In another process in the experiment, a small part of the object is set in the FOV, which is marked with a green square in Fig. 9. When the positions $\Omega_t(x,y)$ of the object at different

times are obtained, they are shown with a white square according to the time series. The center coordinates of those areas are calculated, and the coordinates are arranged according to the time (Fig. 10). A video is created for this real-time tracking process (see Visualization 1). The video is played at 30 fps for a total 20 seconds. In Figs. 9 and 10, the parameter tm denotes one measurement duration, which is equal to approximately 5.6 milliseconds.

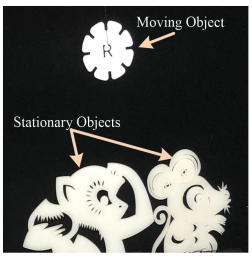

Fig. 11. Complex scene contains two stationary objects and one moving object.

Second, the object is tracked in the complex scene shown in Fig. 11, which contains two stationary objects and one moving object. The moving object is the white button used in the first experiment, and the two stationary objects (monkey and mouse) are fixed in the uniform black scene and serves as the complex background. The moving object is located at a distance of ~1.9m from the imaging system, and size of the illuminated area by the DMD is approximately 15×15cm at the object space. We first illuminate the complex background and detect the reflected light, and then we save the corresponding probe values. Next, the moving object enters the scene and performs a random motion, and then the reflected intensities are received. The background subtraction technique is used to invert the real-time projection curves of the moving object. The results of the projection curves in two directions at different times are shown in Fig. 12, which indicates that the region of the object can be obtained and the corresponding region is set to white. Similarly, 600 sets of experimental results are obtained, and a video is made using those results (see Visualization 2). The video is played at 30 fps for a total 20 seconds. When the area information $\Omega_t(x,y)$ of the object is acquired, the center coordinates of the area are calculated (Fig. 13).

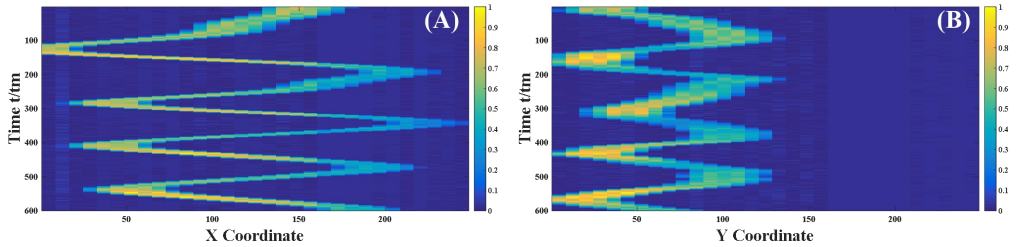

Fig. 12. Experimental results for the projection curves at different times: A) projection curves of the X coordinate and B) projection curves of the Y coordinate. The parameter tm denotes one measurement duration, which is equal to approximately 5.6 milliseconds.

The experimental results show that with an extremely low sampling rate of ~0.20%, real-time tracking of the moving object with 256×256 pixel resolution at ~177 fps can be achieved using the method proposed in this paper. The object tracking accuracy will be affected by the motion blur, which is related to the motion speed, frames-per-second and pixel resolution of the imaging system. When the stroke caused by the motion is less than the pixel resolution during the exposure time of one frame, the tracking accuracy is almost unaffected, and vice versa. Moreover, the frequency of acquiring the position of the moving object will double (~354 fps @ 256×256 pixel resolution) using the complementary modulation acquisition scheme for DMD-SPI [8]. With the development of DMD technology, the modulation frequency will increase and the frequency of the proposed technology will accelerate. The

obtained frequency can reach thousands of fps when the resolution is reduced, which will greatly improve the tracking efficiency of the object. This procedure has the potential for use in real-life applications, such as remote sensing and smart transportation.

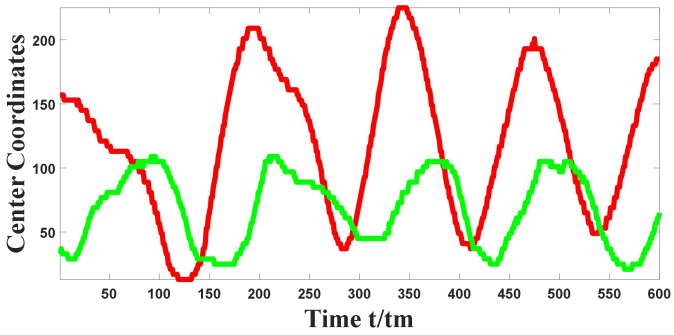

Fig. 13. Center distribution curve. The red line is the x-axis coordinate, and the green line is the y-axis coordinate.

4. Discussion and conclusions

Investigating the ability to track moving objects via SPI is valuable work. In this paper, we proposed a fast object tracking technique based on SPI with an ultra-low sampling rate that is independent of imaging. The modulation information that satisfies the projection conditions is constructed and SPI is able to obtain the projection curves of a moving object in real time and realize real-time tracking of the moving object with high resolution. The number of samples required by this method is extremely low, and background interference can also be overcome by using the background subtraction method. The results of computational simulations and laboratory experiments demonstrate the effectiveness of the proposed method. Currently, our method has three limitations. First, the proposed method can track only one single moving object. Since the method proposed in this paper only uses two projection direction curves, entanglement solutions will be produced when there are multiple moving objects. Second, in a complex background, when the tracked object and background objects have the same reflectivity, the overlap will result in measurement error. Third, the Russian Dolls ordering used in the proposed method will reduce resolution, and result in measurement error. In a future study, pattern recognition and deep learning algorithms will be combined to further overcome those limitations.

Funding

National Natural Science Foundation of China (11404344, 41505019, 41475001), CAS Innovation Fund Project (CXJJ-17S029, CXJJ-17S063) and the Open Research Fund of Key Laboratory of Optical Engineering, Chinese Academy of Sciences (2017LBC007).

References

1. A. Gatti, E. Brambilla, M. Bache, and L. A. Lugiato, "Ghost imaging with thermal light: Comparing entanglement and classical correlation, " Phys. Rev. Lett. **93**(9), 093602 (2004).
2. J. H. Shapiro, "Computational ghost imaging, " Phys. Rev. A **78**(6), 061802 (2008).
3. S. M. M. Khamoushi, Y. Nosrati, and S. H. Tavassoli, "Sinusoidal ghost imaging, " Opt. Lett. **40**(15), 3452-3455 (2015).
4. N. A. Tian, Q. C. Guo, A. L. Wang, D. L. Xu, and L. Fu, "Fluorescence ghost imaging with pseudothermal light, " Opt. Lett. **36**(16), 3302-3304 (2011).
5. O. Katz, Y. Bromberg, and Y. Silberberg, "Compressive ghost imaging, " Appl. Phys. Lett. **95**(13)**,** 131110 (2009).

6. B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. J. Padgett, "3D Computational Imaging with Single-Pixel Detectors, " Science **340**(6134), 844-847 (2013).
7. M. J. Sun, M. P. Edgar, G. M. Gibson, B. Q. Sun, N. Radwell, R. Lamb, and M. J. Padgett, "Single-pixel three-dimensional imaging with time-based depth resolution, " Nat. Commun **7,** 12010 (2016).
8. F. Soldevila, P. Clemente, E. Tajahuerce, N. Uribe-Patarroyo, P. Andres, and J. Lancis, "Computational imaging with a balanced detector," Sci. Rep. **6**(2016).
9. H. Jiang, S. Zhu, H. Zhao, B. Xu, and X. Li, "Adaptive regional single-pixel imaging based on the Fourier slice theorem, " Opt. Express **25**(13), 15118-15130 (2017).
10. C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," Sci. Rep. **8**(2018).
11. Krzysztof M. Czajkowski, Anna Pastuszczak, and Rafał Kotyński, "Real-time single-pixel video imaging with Fourier domain regularization" Opt. Express **26**, 20009-20022 (2018).
12. Z. H. Xu, W. Chen, J. Penuelas, M. Padgett, and M. J. Sun, "1000 fps computational ghost imaging using LED-based structured illumination," Opt. Express **26**, 2427-2434 (2018).
13. J. Huang and D. F. Shi, "Multispectral computational ghost imaging with multiplexed illumination," J. Optics **19**(7)**,** 075701 (2017).
14. D. F. Shi, J. M. Zhang, J. Huang, Y. J. Wang, K. Yuan, K. F. Cao, C. B. Xie, D. Liu, and W. Y. Zhu, "Polarization-multiplexing ghost imaging," Opt. Laser. Eng., **102**, 100-105 (2018).
15. D. F. Shi, S. X. Hu, and Y. J. Wang, "Polarimetric ghost imaging," Opt. Lett. **39**(5), 1231-1234 (2014).
16. P. A. Morris, R. S. Aspden, J. E. C. Bell, R. W. Boyd, and M. J. Padgett, "Imaging with a small number of photons," Nat. Commun **6**(2015).
17. M. P. Edgar, G. M. Gibson, R. W. Bowman, B. Sun, N. Radwell, K. J. Mitchell, S. S. Welsh, and M. J. Padgett, "Simultaneous real-time visible and infrared video with single-pixel detectors," Sci. Rep. **5,** 10669 (2015).
18. M. J. Sun, L. T. Meng, M. P. Edgar, M. J. Padgett, and N. Radwell, "A Russian Dolls ordering of the Hadamard basis for compressive single-pixel imaging," Sci. Rep. **7**, 3464 (2017).
19. E. R. Li, Z. W. Bo, M. L. Chen, W. L. Gong, and S. S. Han, "Ghost imaging of a moving target with an unknown constant speed," Appl. Phys. Lett. **104**(2014).
20. H. Li, J. Xiong, and G. H. Zeng, "Lensless ghost imaging for moving objects," Opt. Eng. **50**(2011).
21. O. S. Magana-Loaiza, G. A. Howland, M. Malik, J. C. Howell, and R. W. Boyd, "Compressive object tracking using entangled photons," Appl. Phys. Lett. **102**(2013).
22. W. K. Yu, X. R. Yao, X. F. Liu, L. Z. Li, and G. J. Zhai, "Compressive moving target tracking with thermal light based on complementary sampling," Appl. Optics **54**, 4249-4254 (2015).
23. G. A. Howland, D. J. Lum, M. R. Ware, and J. C. Howell, "Photon counting compressive depth mapping," Opt. Express **21**, 23822-23837 (2013).
24. R. S. Aspden, N. R. Gemmell, P. A. Morris, D. S. Tasca, L. Mertens, M. G. Tanner, R. A. Kirkwood, A. Ruggeri, A. Tosi, R. W. Boyd, G. S. Buller, R. H. Hadfield, and M. J. Padgett, "Photon-sparse microscopy: visible light imaging using infrared illumination, " Optica **2**(12), 1049-1052 (2015).
25. G. M. Gibson, B. Q. Sun, M. P. Edgar, D. B. Phillips, N. Hempler, G. T. Maker, G. P. A. Malcolm, and M. J. Padgett, "Real-time imaging of methane gas leaks using a single-pixel camera," Opt. Express **25**(4), 2998-3005 (2017).
26. T. Vasile, V. Damian, D. Coltuc, and M. Petrovici, "Single pixel sensing for THz laser beam profiler based on Hadamard Transform," Opt. Laser. Technol. **79**, 173-178 (2016).
27. Z. Zhang, S. Liu, J. Peng, M. Yao, G. Zheng, and J. Zhong, "Simultaneous spatial, spectruml, and 3D compressive imaging via efficient Fourier single-pixel measurements," Optica **5**(3), 315-319 (2018).
28. N. Huynh, E. Zhang, M. Betcke, S. Arridge, P. Beard, and B. Cox, "Single-pixel optical camera for video rate ultrasonic imaging," Optica **3**(1), 26-29 (2016).
29. Y. W. Zhang, M. P. Edgar, B. Q. Sun, N. Radwell, G. M. Gibson, and M. J. Padgett, "3D single-pixel video," J Optics **18**(2016).
30. S. Ota, R. Horisaki, Y. Kawamura, M. Ugawa, I. Sato, K. Hashimoto, R. Kamesawa, K. Setoyama, S. Yamaguchi, K. Fujiu, K. Waki, and H. Noji, "Ghost cytometry," Science **360**, 1246-1251 (2018).
31. Z. B. Zhang, X. Y. Wang, G. A. Zheng, and J. G. Zhong, "Fast Fourier single-pixel imaging via binary illumination," Sci. Rep. **7**(2017).
32. Q. Guo, H. W. Chen, Z. L. Weng, M. H. Chen, S. G. Yang, and S. Z. Xie, "Fast time-lens-based line-scan single-pixel camera with multi-wavelength source," Biomed Opt. Express **6**, 3610-3617 (2015).
33. Y. Jauregui-Sanchez, P. Clemente, P. Latorre-Carmona, E. Tajahuerce, and J. Lancis, "Signal-to-noise ratio of single-pixel cameras based on photodiodes," Appl. Optics **57**, B67-B73 (2018).
34. M. I. Akhlaghi and A. Dogariu, "Tracking hidden objects using stochastic probing," Optica **4**, 447-453 (2017).

35. D. Y. Liu, J. W. Gu, Y. Hitomi, M. Gupta, T. Mitsunaga, and S. K. Nayar, "Efficient Space-Time Sampling with Pixel-Wise Coded Exposure for High-Speed Imaging," IEEE T. Pattern. Anal. 36(2), 248-260 (2014).
36. R. C. Gonzalez and R. E. Woods, Digital Image Processing (Pearson Education, 2011), Chap. 10.